DOCUMENT RESUME

ED 440 456 EA 030 371

AUTHOR	Clune, William
TITLE	Toward a Theory of Systemic Reform: The Case of Nine NSF Statewide Systemic Initiatives. Research Monograph No. 16.
INSTITUTION	National Inst. for Science Education, Madison, WI.
SPONS AGENCY	National Science Foundation, Washington, DC.
PUB DATE	1998-12-00
NOTE	52p.
CONTRACT	RED-9452971
AVAILABLE FROM	Center Document Service, Wisconsin Center for Education Research, 1025 W. Johnson St., Room 242, Madison, WI 53706-1796 ($7). Tel: 608-265-9698 (no phone orders).
PUB TYPE	Reports - Research (143)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Case Studies; *Educational Change; Elementary Secondary Education; *Federal Programs; Government School Relationship; Partnerships in Education; *Program Development; *Program Effectiveness; School Effectiveness
IDENTIFIERS	National Science Foundation; *Statewide Systemic Initiative

ABSTRACT

This paper has two main purposes: (1) to test the central thesis of systemic reform; and (2) to derive lessons about strengths and weaknesses of reform strategies used in policy and practice. It does this by a secondary analysis of case studies of nine Statewide Systemic Initiatives (SSIs) funded by the National Science Foundation. The case studies generated similar kinds of data in categories for all nine systemic reform efforts operating during the same time period, thus permitting a "snapshot" of parallel reforms. The article outlines the central thesis of systemic school reform, which is based on the belief that the increased coherence of instructional guidance is the only way to create large numbers of effective schools. The text examines systemic reform, systemic policy, systemic curriculum, and systemic student achievement. It describes the methodology used in the case studies, rates the nine states, and asks if the SSI Program was successful. The paper details generalizations about the evaluation and offers a profile of successful SSIs. It describes student assessments and teacher networks, discusses missing pieces in the reform landscape, and examines the forces that influence curriculum content. Two appendices offer further statistics on the nine programs and outline protocols for rating systemic reforms. (Contains 17 references.) (RJM)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 440 456

National Institute for Science Education

University of Wisconsin-Madison • National Center for Improving Science Education

Research Monograph No. 16

Toward a Theory of Systemic Reform: The Case of Nine NSF Statewide Systemic Initiatives

William Clune

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

EA 030371

Funded by the
National Science Foundation

BEST COPY AVAILABLE

National Institute for Science Education (NISE) Publications

The NISE issues papers to facilitate the exchange of ideas among the research and development community in science, mathematics, engineering, and technology (SMET) education and leading reformers of SMET education as found in schools, universities, and professional organizations across the country. The NISE Occasional Papers provide comment and analysis on current issues in SMET education including SMET innovations and practices. The papers in the NISE Research Monograph series report findings of original research. The NISE Conference and Workshop Reports result from conferences, forums, and workshops sponsored by the NISE. In addition to these three publication series, the NISE publishes Briefs on a variety of SMET issues.

The research reported in this paper was supported by a cooperative agreement between the National Science Foundation and the University of Wisconsin–Madison (Cooperative Agreement No. RED-9452971). At UW–Madison, the National Institute for Science Education is housed in the Wisconsin Center for Education Research and is a collaborative effort of the College of Agricultural and Life Sciences, the School of Education, the College of Engineering, and the College of Letters and Science. The collaborative effort is also joined by the National Center for Improving Science Education, Washington, DC. Any opinions, findings, or conclusions are those of the author and do not necessarily reflect the view of the supporting agencies.

Research Monograph No. 16

Toward a Theory of Systemic Reform: The Case of Nine NSF Statewide Systemic Initiatives

William Clune

National Institute for Science Education
University of Wisconsin–Madison

December 1998

About the Author

William H. Clune is Voss-Bascom Professor of Law at the University of Wisconsin Law School, co-team leader of NISE's Systemic Reform: Policy and Evaluation Team, and a senior researcher of the Consortium for Policy Research in Education. His past research has included school finance, school law, implementation, special education, public employee interest arbitration, school site autonomy, effects of high school graduation requirements, upgrading of the high school curriculum in mathematics and science, and systemic educational policy. His present research includes "program adequacy" (the cost and implementation structure needed to reach high minimum levels of student achievement in low-income schools), systemic policy in mathematics and science education, and systemic reform in urban districts.

Acknowledgments

Thanks to the following people for reading and commenting on earlier drafts: Bernice Anderson, Daryl Chubin, Eric Hamilton, Eric Osthoff, Andy Porter, Senta Raizen, Patrick Shields, Debbie Stewart, Larry Suter, Norman Webb, Paula White, and Andy Zucker. The comments were helpful; errors are mine.

Contents

Figures and Tables v

The Central Thesis of Systemic Reform 1

Testing the Central Thesis: A Theory of Systemic Policy and Reform 2

Operationalizing the Variables 3

Systemic Reform 3

Systemic Policy 3

Systemic Curriculum 4

Systemic Student Achievement 4

Methodology 5

The SRI Case Studies 5

Limits from Studying the NSF Initiatives, Including Measuring Partial Causation 5

Results

Rating the States 7

Was the SSI Program Successful? 9

The Imprecise Task of Testing Causation 9

Generalizations and Cross-Cutting Themes 10

The Typical Profile of Successful SSIs 10

The Importance of Earlier Periods of Reform and the Time Required for Successful Reform 10

Student Assessments and Teacher Networks as the Universal Middle Link 11

Limits of the Sequential Causal Theory: "Systemic Causation" in Mature Cultures of Reform 11

Some Missing Pieces in the Reform Landscape 11

The Absence or Indirection of Influence over Curriculum Content 11

The Dearth of Fully Aligned State Assessments 12

The Absence of Good Data and Evaluation of the Impacts of Reform on Classrooms and Student Achievement 12

The Slow Growth of Incentives and Mechanisms for Whole-School Restructuring .. 13

The Unexplored Territory of Adequacy and Cultural Context in Urban Schools 13

Conclusion: Making a Difference Using Theory to Build New Reform 13

References 15

Appendixes

A: Nine Statewide Systemic Initiatives Studied in SRI Case Studies, Synopsized and Ranked According to the NISE Theory 17

B: NISE Protocol for Rating Systemic Reforms 25

Figures and Tables

Figure 1. Model of systemic reform used to guide SRI case studies 6

Table 1. Breadth, depth and average ratings of the 9 SRI states 8

This paper has two main purposes: (1) testing the central thesis of systemic reform and (2) deriving lessons about strengths and weaknesses of actual reform strategies that are used in policy and practice. Both purposes will be pursued through secondary analysis of a convenient source of data, case studies (SRI, 1998) of nine Statewide Systemic Initiatives (SSIs) funded by the National Science Foundation (see references; summary of case studies in Appendix A). The case studies collect similar kinds of data in useful categories for all nine systemic reform efforts operating during the same time period (1992–96), thus permitting a methodologically controlled "snapshot" of parallel reforms. The case studies of SSIs also allow the sponsor of this paper, the National Institute for Science Education (also NSF funded), to learn from the experience of the SSIs in its study of systemic reform.

The Central Thesis of Systemic Reform

As framed by Smith and O'Day (1991), the central thesis of systemic school reform is that greater coherence (or alignment) of policies of instructional guidance (those affecting the content and quality of instruction in schools) is the only way to create large numbers of effective schools (schools producing desirably high levels of student achievement). The specific kinds of policies mentioned in their model have persisted as the assumed components of systemic reform: curriculum frameworks, instructional materials and curricula, inservice professional development, preservice professional development, student assessments and accountability, school site autonomy and restructuring, and supportive services from districts and the state. Although "policy" at the top was seen as the driving force for change, systemic reform was not defined exclusively in top-down terms. Inservice professional development was seen as depending on active networks of teachers organized from the grass roots. School restructuring is another feature that might be stimulated by government action, but obviously could not occur at that level. Indeed, systemic reform was proposed by Smith and O'Day partly as a way of generalizing (or going to scale with) successful models of school restructuring developed during a prior period of decentralized reform.

Smith and O'Day posited another element of systemic reform: standards-based curricula as the touchstone for policy alignment, modeled on the pioneering standards for mathematics developed by the National Council of Teachers of Mathematics (NCTM). Standards-based curricula aim for active learning by students and support teaching for understanding (Cohen, D., McLaughlin, M., & Talbert, J., 1993; Mintzes, J., Wandersee, J., & Novak, J., 1998), as opposed to the exclusive emphasis on basic skills that characterized some earlier (and probably somewhat successful) exercises of policy alignment, such as minimum competency achievement testing. Both the meaning of teaching for understanding and the proper emphasis to be placed on basic skills are hotly debated to this day (Standards, 1998; Math articles, 1999). But some kind of deepening (or upgrading) of the curriculum has remained a universally accepted goal of systemic reform, especially for disadvantaged students. Thus, the terms *systemic reform* and *standards-based reform* have become virtually synonymous (Knapp, 1997). While this paper does examine systemic reform at the state level, Smith and O'Day's exclusive focus on the state as the locus of policy has been broadened. Many large districts are creating their own systems, and NSF now has a program funding Urban Systemic Initiatives.

Building on this background, we can state the *central thesis of systemic reform* as follows:

> Systemic reformers can bring about a greater degree of alignment of policies of instructional guidance around new standards of learning, thereby producing widespread and substantial gains in the quality of teaching and learning for all students throughout the area affected by the policies.

Testing the Central Thesis: A Theory of Systemic Policy and Reform

In order to test the central thesis, we needed to develop a testable theory. The theory presented in this paper follows the central thesis but also reflects the practice of the NSF-funded SSIs. Three researchers, William Clune, Eric Osthoff, and Paula White, gathered data about all of the SSIs from workshops, forums, and interviews with systemic reformers and researchers, as well as from documents, such as proposals and evaluations. A book manuscript applying the theory will be written later based on the broader set of data. While this paper applies the theory to the data set made available by the nine SRI case studies, all three members of the research team found no major inconsistencies between the two studies, except that, in the larger study, the ratings of SSIs (including the nine SSIs common to both) may be lower and the findings about successful models more varied, both across states and across reform components within states. Comparing our theory and findings with SRI's provides an additional checkpoint on validity and usefulness, and we welcome feedback that might further shape the proposed book.

A good theory of systemic reform should model the indispensable elements of the central thesis of systemic reform: a *policy system* (including an unspecified mix of policies and intermediate organizations and activities) with a strong influence on *a rigorous curriculum as actually taught to all* students (though possibly a differentiated curriculum) and corresponding measured *high student performance* and *systemic reform*: some set of activities that bring systemic policy into existence. These basic elements, shown schematically in causal relationship, look like this:

> *Systemic reform* (SR), through its purposeful activities, leads to
> *Systemic policy* (SP), which leads to
> *A rigorous implemented curriculum* (SC) for all students, which leads to
> Measured *high student achievement* (SA) in the curriculum as taught

This kind of system is dynamic even in its fully mature state (requiring constant communication and adaptation), and even successful reform will likely proceed incrementally (with more reform leading to gradually stronger policies, leading to gradually stronger curriculum for more students and greater gains in student achievement), so that systemic reform obviously should be represented as a *continuous causal sequence*:

$$SR \rightarrow SP \rightarrow SC \rightarrow SA$$

where SR = systemic reform, SP = systemic policy, SC = systemic curriculum, and SA = student achievement corresponding to the curriculum.

Operationalizing the Variables

To test the above model against real reform efforts requires three things beyond the schematic: first, the variables must be made specific and measurable (operationalized); second, they must be operationalized in a way that corresponds to the causal theory; and third, the measurement must show to what extent the goal has been achieved of changing the entire system rather than a few teachers, schools, or students.

We decided to meet all three requirements by conceptualizing the variables according to characteristics or elements that make them influential and then rating overall variables (taking all elements into account) on five-point scales of breadth and depth. Breadth in our method refers to the scope of the variable across the elements, and a score of 5 would be given if all the elements were present. Depth refers to the strength of the influence, combined with its quality, or adherence to the model of standards-based reform, with a score of 5 being awarded for maximum quality and strength. Appendix B of this paper is a detailed matrix that displays our rating system by variable, component of each variable, and criteria for rating the breadth and depth of each component (Eric Osthoff prepared the matrix for the larger project). A narrative summary of that matrix is given below.

Systemic reform. After studying data on all the SSIs, we decided to conceptualize systemic reform as "reform leadership and management." The influence of this variable in any state involves the following elements: vision, strategic planning, networking with policymakers, networking with professionals, institutionalization of the reform structure, leveraging of resources, and public outreach and visibility. The reform would be considered broad to the extent it had all of these elements, and the elements touched all the levers of policy, and deep to the extent that each element was strong and of high quality, defined as conforming to a standards-based vision of reform.

Systemic policy. The components of the policy system that are rated for breadth and depth are curriculum standards; curriculum frameworks; student assessments; instructional materials; equity targeting policies; preparation and initial licensing of teachers; teacher recertification; professional development for teachers and administrators; accountability for students, teachers, schools, and administrators; and district and school capacity-building and improvement. The policy system would be considered broad to the extent that it covered the full range of influential policies in the area, and that the policies themselves covered the full range of subjects, grades, and schools; it would be considered deep to the extent that it has strong predicted influence on schools, teachers, and students and pushed in the direction of standards-based teaching and learning. We decided to conceptualize the strength of the policy components according to a set of attributes developed by Porter and colleagues for this very purpose (Porter, Floden, Freeman, Schmidt, & Schwille, 1988). Details on measuring strength of policy are given in the next paragraph.

Systemic policy's strength (influence) is defined by the strength of four attributes—authority, power, consistency, and prescriptiveness or detailed guidance—each of which can be reflected in a variety of specific policies and organizational forms, depending on the context. *Authority* is provided through the backing of powerful institutions and individuals, such as the governor,

legislature, or intermediate network of teachers or professional organizations. Sometimes a particular policy instrument, such as student assessment, achieves a kind of authoritative recognition. Some states, particularly in the South, seem to have governmental authority structures that are especially well accepted in districts and schools. *Power* is attained through resources, such as professional development opportunities or financial rewards, or through other incentives, such as the stakes attached to a student assessment or an accountability system. *Consistency* is the extent to which all the elements of influence push in the same direction and are aligned around a common vision and content. *Prescriptiveness*, or detailed guidance, is the extent to which the policy system gives a clear idea of exactly what schools and teachers are supposed to do through, for example, the availability of textbooks, replacement curriculum units, student assessments, and demonstration teaching tapes.

Systemic curriculum. Content and pedagogy, the material actually conveyed to students in classrooms and the instructional methods by which it is taught, make up systemic curriculum. Content refers to the knowledge or skill that students are supposed to learn in subject areas like algebra and geometry, as well as skill areas like computation, problem solving, and conceptual understanding. Pedagogy refers to the kind of teaching that is employed, particularly whether the demands on students match the content and skills that are being taught, for example, whether students actually solve and discuss problems if the goals are problem solving and communication. Breadth depends on the number of schools, teachers, grades, subjects (math, science, etc.) that demonstrate change. Depth depends on the extent of the change. Deep change would refer to substantial upgrading of the content and a correspondingly strong change in pedagogy. Shallow change refers to smatterings or layerings of new content and pedagogy, a common finding for the extent of curriculum reform and perhaps its greatest challenge (Knapp, 1997). Also pluses for curriculum breadth and depth are equity targeting in the curriculum and school improvement aimed at curriculum change. We also considered the availability of good data on curriculum as part of its depth because good data help guide reform. But the availability of data would inevitably be reflected in the depth rating in any case because good data are helpful in showing deep curriculum change. As explained below, systematic observational data on the implemented curriculum were rare, but were considered a definite plus where they occurred (teacher surveys and observations at selected sites were common; other indicators are more indirect, such as whole-school curriculum reform).

Systemic student achievement. The primary measure of systemic student achievement is gain on a student assessment in some way aligned with the reform (for example, gain after stronger policies were enacted or gain in schools receiving more emphasis under the policies). Assessments commonly available in the states with SSIs included state assessments and NAEP. Some state assessments are better aligned with the goals of policy than others (a fact that would be reflected in the consistency rating of the policies). Gains in equity (gap closing) were counted as a plus, as were gains in course enrollment and attainment in later grades. Breadth of gain in student achievement depended, once again, on how many students, schools, grades, and subjects showed gains. Depth refers to the size of the gains, as well as the quality of the data on achievement. A gain over five years of one or two percentage points of the total number of students in the state reaching proficiency on a student assessment seems relatively small in terms of policy goals and was at the small end of our sample (1 on a scale of 5), while a gain of 8 or more points seems large and was at the high end of the sample (5 on a scale of 5).

Methodology

The methodology for this paper consisted of, first, reading and taking detailed notes on all nine of the SRI case studies of SSIs (1998), focusing on what appeared to be strong evidence related to our theoretical categories; and second, rating all the variables in every state on both depth and breadth according to our theoretical model and the rating matrix previously discussed (see Appendix B).

The SRI Case Studies

For the 1998 case studies from which this paper was drawn, SRI used a model that depicted "SSI Activities" as affecting a foundation of policy, which in turn affected teachers, schools, and student achievement. The graphic form of the model is given in Figure 1.

Following this model, the case study researchers gathered data in each of the nine states according to implementation of the reform, effects on policy, effects on teachers (meaning effects on how teachers were trained and taught), and effects on students. These categories nicely fit the four main variables in our model of systemic reform and management, systemic policy, systemic curriculum, and systemic student achievement.

The information in the SRI reports was translated into the theoretical framework used in this paper in two steps. Appendix A gives our narrative synopsis of each SSI by each of the four main variables (reform, policy, curriculum, achievement) and, in addition, includes a general comment on the overall strength of the reform. Appendix A is long and detailed, but readers unfamiliar with the data, and looking for the human (or at least organizational) face of reform, should find it very helpful as a way of grounding the analysis. We used the narrative synopsis to develop a numerical rating of every element of every reform in both breadth and depth. The results of that rating are given in the section on Results.

Limits from Studying the NSF Initiatives, Including Measuring Partial Causation

A number of readers of earlier drafts of this paper asked whether our theory is of systemic reform generally or only of the NSF-funded Statewide Systemic Initiatives. The short answer is that we see many of the SSIs as good examples of systemic reform and the whole group as a good test-bed for the theory, but we concede that some limitations and complexities of analysis flow from our focus on the NSF SSIs. The guidelines issued by NSF for proposals from the states reflected the Smith and O'Day formulation, and most states built their reforms roughly along those lines. It is true that some reforms focused heavily on professional development funded by the SSI itself (reflected in the SRI graphic by the arrow running directly from the SSI to the schools); another approach was pilot schools combined with varying degrees of emphasis on policy. Regardless of the actual approach of the reform, we tested the prediction in our model that reform could change schools only through increasing policy alignment. Thus, a state that produced big changes in curriculum and achievement without affecting policy (solely through its own professional development activities, for example) would be counted as evidence against the validity of our theory. In other words, our model predicts that success of the SSIs in changing

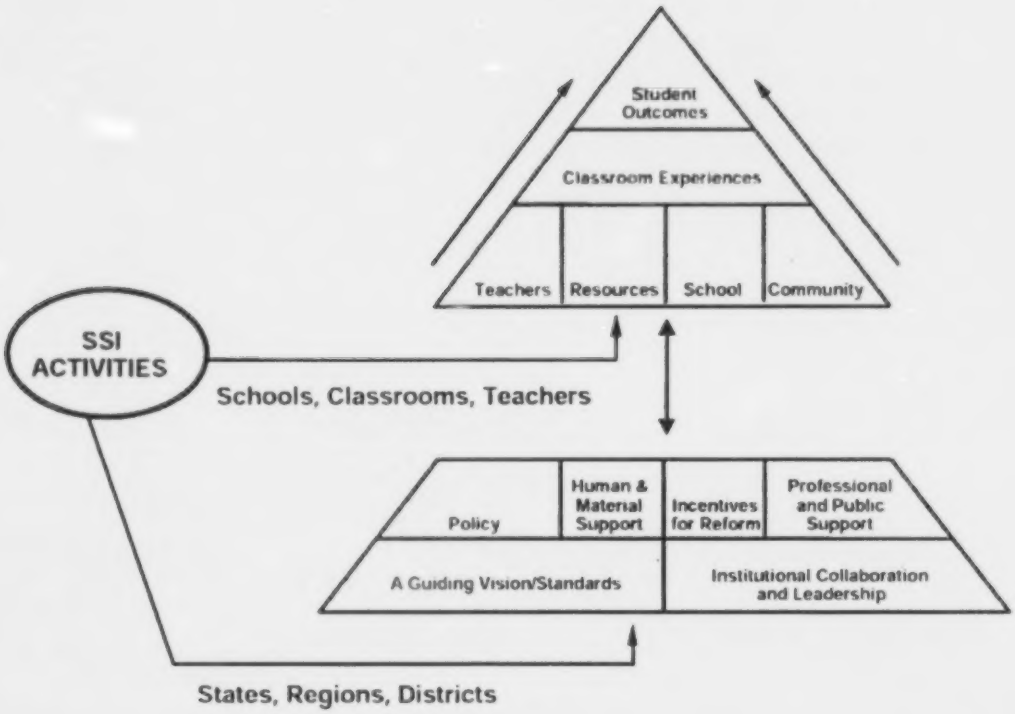

Figure 1. *Model of Systemic Reform Used to Guide SRI Case Studies*

Source: Zucker, A. A., Shields, P. M., Adelman, N. E., Corcoran, T. B., & Goertz, M. E. (1998). *A Report on the Evaluation of the National Science Foundation's Statewide Systemic Initiatives (SSI) Program.* Menlo Park, CA: SRI International.

schools will be determined by how closely they follow the classic model of systemic reform, and the states taking a different approach provide us with needed comparison strategies.

A second and related complexity is the relationship of the SSIs to other systemic reforms, both in the same state and in other states not studied or funded. For related reforms in the same state, we had to judge (as did the SRI researchers) whether the SSI made a substantial contribution to the increased degree of alignment, if any. I noticed a similar sense of partial causality in a statement on the Weather Channel, "The above average number of storm-related deaths in California this summer was undoubtedly due in part to El Niño." If we wanted to carry this analogy out, El Niño would correspond to the NSF-funded systemic reformers, the storms would correspond to policy alignment, the swollen rivers would correspond to an upgraded curriculum, and the storm-related deaths would correspond with student achievement. Given this model of partial causation, other reforms occurring in the state at the same time might also get credit for pushing toward systemic

policy, and, indeed, we found that the prior enactment of a standards-based student assessment was an important stimulus to reform. Another limiting effect of the focus on the SSIs is that we have no data on states that did not receive any NSF funding. From our data base we do not know whether other states achieved equal levels of systemic reform without such funding.

The issue of partial causation and how to recognize it deserves further discussion, because it operates at every stage of our model. Systemic reforms join other forces in leading to stronger policies. Stronger state policies may not be the only cause of curriculum improvements (higher course requirements from an earlier time being another); and curriculum improvements may not be the sole cause of increases in student achievement (demographic changes being another candidate; for example, unmeasured, gradual increases in higher education among parents). We (and I think it is fair to say the SRI researchers) took two approaches to the recognition and measurement of partial causation: qualitative and quantitative. Qualitatively, we looked for anecdotal evidence that the activity at one stage of the model was being felt at the next stage, for example, that a curriculum designed by the reformers actually was adopted in policy, adopted by schools, and reflected in student achievement. Gains in student achievement that did not seem associated with the presence of reform in schools, or that occurred in a time period too early to reflect the impact of reform, would be assumed the result of some other factor. Quantitatively, once we had some confidence in the basic correspondence between activities in each stage of the model, we would then measure the breadth and depth of those changes and see whether high ratings at one stage corresponded to high ratings in the next. This methodology for measuring partial causation is fuzzy and inexact, but seems reasonably robust in practice. Reforms have a logic of action that can be plumbed by careful evaluation, as in the SRI case studies.

A third complexity is how much the limited time period analyzed in the case studies can tell us about the progress of reform over a longer period of time (especially since one of our findings is that the more successful reforms built on past reforms and typically were incomplete at the end of five years). The answer is that the case studies must be considered a "snapshot" of reform in progress over the five years. If a reform had reached the stage of greater alignments in policy, but had not reached many schools, it would get high ratings on policy but lower ratings on curriculum and achievement. As will be seen, Louisiana turned out to be a state where student achievement had not yet responded strongly to reform. A different kind of case is where the reform strategy adopted in the first five years was judged ineffective and dropped in favor of a more promising strategy. A reform that was "just getting its act together" at the end of the first five-year period (as actually occurred in some cases) would get a low rating then, based on the NISE system, and would deserve that rating, but might get a high rating using the same criteria at a later point in time.

Results

Rating the States

The results of the rating exercise are given below in Table 1, with the states listed from highest to lowest in the average of all ratings.

Table 1. Breadth, depth and average ratings of the 9 SRI states

STATE	REFORM		POLICY		CURRIC.		ACHIEVE.		STATE AVG.
	Br.	Dp.	Br.	Dp.	Br.	Dp.	Br.	Dp.	
Connecticut	4	4	4	4	3	2	4	4	3.6
Maine	4	4	4	4	3	2	4	4	3.6
Montana	3	4	2	4	2	3	2	4	3.0
Louisiana	4	4	3	2	3	2	2	2	2.8
Michigan	2	3	2	2	2	2	3	2	2.3
California	2	3	2	3	3	2	2	1	2.3
Arkansas	3	3	2	2	2	1	2	2	2.1
Delaware	2	1	1	1	1	1	1	1	1.1
New York	1	1	1	1	1	1	1	1	1.0
AVERAGE OVER STATES	2.8	3.0	2.3	2.6	2.2	1.8	2.3	2.3	

Let's begin discussion with what can be concluded from this quantitative analysis. The question of whether strong Reforms led to stronger Policy, which led to a stronger Curriculum, which led to stronger Achievement can be assessed by reading backward from Achievement. Higher ratings on Achievement are associated with higher ratings for the other variables, particularly Reform and Policy. Additional support for these relationships comes from the correspondence of our ratings with the funding renewal decisions of NSF (those decisions themselves emerging from careful performance reviews and ratings by panels of expert reviewers). Two of the top three states in Table 1 had their funding renewed by NSF (Connecticut, Maine). The third state getting renewed funding, Louisiana, ranked fourth in our analysis and had an average rating of 2.8, slightly behind Montana. Almost surely (but not entirely without controversy), Montana was downgraded because of its exclusive emphasis on high school mathematics, reflected in lower breadth ratings for our variables. As for Louisiana, some commentators have suggested that equity (percent of minority students) may have played a role in the refunding of this state. Equity is an announced goal of the SSI program (success for all students) and would be a legitimate basis for decision in a close case. Another compatible explanation is that Louisiana has the same high ratings for Reform as Maine and Connecticut. The strength of the reform base makes likely a strong future impact on Curriculum and Achievement.

A second set of observations can be made about the variables looking at the average ratings of variables over states in the bottom row. Reform and Policy are stronger than Curriculum and Achievement, and, within Reform and Policy, depth (or strength of influence) is stronger than breadth (coverage of the whole state). Greater strength in Reform and Policy can be expected because of both the sequence of reform (with those areas receiving attention first) and the sheer difficulty of making an impact on teachers and students. Greater depth than breadth of reform might be expected because reformers will discover strong reform and policy tools before extending them to the whole system. The generally lower ratings for Curriculum and Achievement reflect some problems of policy design, plus major problems of data and

measurement. Both of these issues are discussed further below. Examples of design problems are the lack of emphasis on curriculum content and whole-school restructuring. The lowest average rating across states is for depth of influence in the curriculum, and exactly this — shallow influence on the curriculum — was identified as the chief failing of systemic reform in an earlier research synthesis sponsored by NISE (Knapp, 1997). The main data problems with Curriculum were scant data and indirect measurement of what was going in classrooms. The main data problem with Achievement was the lack of alignment of student assessments with the goals of reform, but the absence of good control groups for evaluation was a close second as a problem.

Was the SSI Program Successful?

What can the ratings of the states tell us about the success of the NSF's SSI program? The highest possible standard of evaluation would be deep and broad change in every aspect of the system in every state. That standard, which would translate into an average of five for every state, was not met. Looking at the last column, the states averaged from just below 4 to 1 on a five-point scale. Even the higher rated states reached at most 50% of this "whole system" target and did so with inconsistent depth and quality. But the standard of perfection is surely too high, given the limited time and resources available to the reformers, the complexity of the systems, and the highly experimental nature of the reforms themselves. A more reasonable standard is whether substantial change occurred in most states, and that standard was met. Only New York and Delaware made no progress, and both of these states were retooling in promising directions at the end of the five-year period. Thus, the reforms seem cost-effective if not massively effective. A good argument for this point of view was given by Zucker and Marder in the case study of Montana, which said that the strategy of "concentration" producing deep change in one sector at a time in some strategic order may be as good an investment of resources as the "holistic" strategies of many states that produced broader but shallower change, (SRI, 1998).

The Imprecise Task of Testing Causation

The primary indicator of causation, a correlation of all the systemic variables, was satisfied as well as could be expected in a sample of nine states. Higher ratings go with higher ratings across all four variables. This rough correspondence should not be understood as anything like a rigorous statistical test. There were only nine cases, with the bulk of the measurements falling closely together in the middle ratings. Differences were small given the sample size and the imprecision of the measurements, as in data on Curriculum and Achievement. And some qualitative judgments were made to derive the numerical ratings. For example, two states, Michigan and Arkansas, showed higher gains on statewide tests of student achievement than are reflected in their ratings. The reason for lower ratings of student achievement in both cases is that the gains shown were judged probably related to an earlier period of basic skills reform, a judgment supported by the intensity of the earlier period of reform, the timing of the gains in achievement, and the lack of any evidence that instruction changed, such as comparisons among units affected to a greater or lesser degree by the policy changes brought about by the SSI.

Generalizations and Cross-cutting Themes

Discovering common patterns of organization and strategy across SSIs requires qualitative analysis, and for this the case study is indispensable. As background for the generalizations about reform that are discussed here, readers are again urged to read through all of Appendix A (qualitative syntheses of each case study organized by the four variables in our model).

The Typical Profile of Successful SSIs

The typical profile in the higher rated states, described according to the four variables of the theory presented in this paper, looks like this: (1) Reform. A reform agency with independence but strong connections with the scientific disciplines in higher education; strong networking of reformers with supportive professional leadership organizations in the state; a mission including both math and science; long–term support of key policymakers, especially the governor. See Treisman (1997) regarding the idea of "working in the middle" as a genotype for successful systemic reform. (2) Policy. A state assessment as a key building block of policy; intensive (cumulating at least four weeks per year) professional development aligned with standards reaching a substantial number of the state's teachers; development of teacher networking built around curriculum and instruction (usually involving both face-to-face and electronic contacts); a workable approach to school improvement; strong connections with preservice teacher education departments in the state universities. (3) Curriculum. A substantial but not transformative influence on curriculum and teaching in the direction of the new standards. (4) Student Achievement. A substantial positive impact on student achievement, something like 10 points on a 100-point scale over 5 years (an average of 2 points per year).

This description of success also fits the lower rated states, where one or more important pieces of the composite picture are missing. In fact, some lower rated states are decisively stronger on selected components of variables. California's teacher networks, for example, probably were the model of design and impact, but political and policy support in that state disintegrated near the end of the initiative. Montana's strategy of curriculum replacement had the greatest impact on the classroom, but the scope of the initiative was limited to high school mathematics.

The Importance of Earlier Periods of Reform and the Time Required for Successful Reform

A pattern that emerges in this group of case studies is that successful states built on pre-existing reforms of the 1980s, with continuity rather than discontinuity between the earlier period and the new period of systemic reform. Usually the first piece was the state assessment itself, which acquired a base of statewide authority and acceptance strong enough to support subsequent modifications in a more standards-based direction. In Montana, the foundation was prior development of a standards-based curriculum and teacher enhancement projects, which then acquired the support of state policy. In any case, the lesson is that reform takes more time than the five years allowed in one cycle of NSF funding.

Student Assessments and Teacher Networks as the Universal Middle Link

The combination of a state assessment as the lead policy instrument and professional networking as a delivery structure operates as a kind of universal link between the top and bottom, regardless of whether state policy is built on central or local control. States with strong centralized policies need a way to bridge the gap between the top and the bottom, while local control states find that the assessment/network format is a politically acceptable way to provide strong instructional guidance. In both kinds of states, assessments and networking bridge the gap between the large "grain size" of the standards and the more specific tasks demanded by teaching and learning (see Standards, 1998).

Limits of the Sequential Causal Theory: "Systemic Causation" in Mature Cultures of Reform

The notion of reform becoming embedded in a student assessment, which in turn becomes embedded in the discourse of a network of teachers, points to a limitation of the sequential causal theory presented in this article. Once teachers are in the "net," they become part of all the "boxes" or variables of reform: reformers, policymakers, curriculum implementers, and facilitators of student achievement. They are reformers and policymakers because they help construct each modification of standards and assessments, and they implement the curriculum and shape student achievement in their own classrooms. Subgroups of teachers take the lead in developing the examinations, working with teachers from higher education, while others focus more exclusively on their own classrooms. To some extent the entire system becomes a "learning organization," in which the causal processes of reform are distributed across roles (Resnick, 1997). This kind of causation in mature systems might be called "systemic causation." Some dispersed causation can be captured within the confines of the NISE model used in this paper, which is labeled a "continuous causal sequence" and whose notion of "depth" does include deeper understanding by all system actors and even cultural change. Further, the multiple roles of teachers can be thought of as adding authority to the policy system. But at some point a system of simultaneous, multidirectional communication requires a more elaborate model (for earlier, less linear modeling, see Clune 1993a, 1993b).

Some Missing Pieces in the Reform Landscape

The previous section dealt with commonalities observed across successful reforms, but the interstate overview provided by the case studies also reveals a number of glaring deficits, or missing pieces, in the reform landscape.

The Absence or Indirection of Influence over Curriculum Content

Although it is true that student assessments and teacher networks served as the link between top and bottom in the reforms, that link would have been stronger with a more powerful means of influencing curriculum. The common problem is the focus on pedagogy rather than content. Reforms typically were aimed at classroom processes such as the use of manipulatives, collaborative learning, and inquiry learning. Especially early in the reforms, direct means of influencing curriculum such as model curricula, new materials, and model teaching units were relatively rare.

Criticism of the pedagogical orientation could easily be overdrawn. Not only is active learning supported as effective by research from cognitive psychology, the distinction between content and pedagogy is not entirely clear. Well-conceived active learning techniques raise the level of cognitive demand or complexity in any given domain of content. Graphing, for example, is not simply a technique of representing a function but a different kind of content and a means of seeing more deeply into the material. Furthermore, many teacher training programs incorporated content as part of the training when, for example, inquiry-based science in elementary school required restructuring the curriculum or curriculum units were used as part of teacher training. Nevertheless, it is surprising how few reforms focused on *what* the students were being taught as opposed to only *how*. The gap between pedagogy and content narrowed as the reforms progressed, partly as a result of productive prodding by NSF. By the mid-1990s, many of the stronger reforms were using new materials, model teaching units, or curriculum replacement units (see Cohen & Hill, 1998, and Kennedy, 1998, for research showing that professional development is more effective when it focuses on content).

The Dearth of Fully Aligned State Assessments

Despite the importance of student assessments in reform, the absence of assessments that are aligned, or fully aligned, with the reform objectives is a constant source of frustration. Reform objectives are neither advanced nor well measured by mismatched assessments. It is true that progress was made during the 1990s as new assessments were developed, piloted, and implemented. And, even in the absence of a fully aligned assessment, a major contribution to testing causal influence could be made by a more detailed understanding of which items on various state assessments are more and less matched to the objectives of reform.

The Absence of Good Data and Evaluation of the Impacts of Reform on Classrooms and Student Achievement

The impact of systemic reform can be recognized without the strongest data on changes in classroom practice and student achievement, but good data and design around these variables would lend considerably more confidence to such judgments. Any theory or evaluation of systemic reform requires testing causal links in complex systems on the basis of relatively few cases (observations). The task would be much easier and the case much more convincing if there were more direct and precise data on teaching and learning that could be associated with varying degrees and phases of reform. States were certainly moving in that direction with, for example, evaluations that compared gains in student achievement with the number of SSI-trained teachers in schools; but the effort is truly in its infancy. In one sense, no excuse exists for not gathering better data on teaching and learning, because adding the measurements is relatively easy and inexpensive compared to the daunting task of changing systems. True, the difficulty of measurement can be underestimated. Measurement of instruction, for example, must include not only pedagogical techniques like active learning but also the rigor and importance of the math or science concepts being taught, appropriate sequencing and connections, and articulation without unnecessary repetition between grades and levels of schooling (thanks to input from Senta Raizen on this point). But the biggest challenge is not in the difficulty but in the timing. The hard part is building good measurement and evaluation design into a program that is being invented and implemented on the fly and always has more urgent priorities. Fortunately, a funding agency

is well equipped to insist on a solution to this problem of timing and priority, and improvement of evaluation should be and has become a major priority in the systemic reform program of NSF.

The Slow Growth of Incentives and Mechanisms for Whole-School Restructuring

Another "late bloomer" on the reform landscape was building incentives for whole school restructuring. Many reforms were better at going to scale with the training of teachers within schools than changing the schools (and districts) in which the teachers would operate, and school restructuring proved a serious obstacle to change. Gradually, components aimed at school restructuring, such as administrative outreach and workshops, became more common. At least one SSI not reviewed in this paper has a powerful model of school restructuring (Rodriguez, in press). But this component appears sufficiently underdeveloped even at this time that it deserves further cross-site study as the basis for better technical assistance to the reforms.

The Unexplored Territory of Adequacy and Cultural Context in Urban Schools

One problem that appeared in such a fragmentary way that it is barely on the radar is the adequacy, or instructional capacity, of urban schools and districts. This problem requires further study to understand its basic dimensions. In the urban areas in some states, the obstacle is shortages of key resources, such as textbooks, materials, and computers. In others, materials are plentiful, but special problems of training exist, due to, for example, rapid turnover. In still others, the obstacle identified is a complex and resistant urban school bureaucracy. Another challenge is making the new curricula accessible in the ethnically pluralistic urban context (Lee, 1998). Finally, student mobility may raise special problems for an articulated multiyear course of instruction and associated data systems on instruction and achievement. The special obstacles to reform in urban districts, as well as, perhaps, the special advantages, deserve further study. Some research already exists (St. John, Century, Tibbits, & Heenan, 1994), and the rapid expansion of the Urban Systemic Initiatives offers an opportunity to look more deeply.

Conclusion: Making a Difference Using Theory to Build New Reform

In this paper, I discussed how a particular theory of systemic reform can be used to conceptually simplify, describe, evaluate, and draw conclusions from case studies of reforms in different states. But the theory also has prospective and practical applications. Every component that is important to success in other reforms can become part of the design of new ones, for example, the independence of the reform agency and its connections with policymakers, teachers, and schools. The historically most powerful tools of policy, such as student assessments and teacher networks, can be raised in priority. Deficits found in earlier reforms can be addressed at the beginning of new ones, such as influence over curriculum content, assessments or items on assessments aligned with reform objectives, whole school restructuring, and good evaluation design. Indeed, it is quite clear from reading the case studies that there has been a learning curve in the systemic reform movement nationally that is a by-product of lessons learned in individual states. Hopefully, the theory offered here can help strengthen that learning process in the future.

References

Clune, W. H. (1993a). The best path to systemic educational policy: Standard/centralized or differentiated/decentralized? *Educational Evaluation & Policy Analysis, 15*, 233-254.

Clune, W. H. (1993b). Systemic educational policy: A conceptual framework. In S. H. Fuhrman (Ed.), *Designing coherent education policy* (pp. 125-140). San Francisco: Jossey-Bass.

Cohen, D. K., & Hill, H. C. (1998). *Instructional policy and classroom performance: The mathematics reform in California* (Research Report No. RR-39). Philadelphia: University of Pennsylvania, Consortium for Policy Research in Education.

Cohen, D., McLaughlin, M., & Talbert, J. (Eds.). (1993). *Teaching for understanding: Challenges for policy and practice*. San Francisco: Jossey-Bass.

Kennedy, M. (1998). *Form and substance in inservice teacher education* (Research Monograph No. 13). Madison: University of Wisconsin–Madison, National Institute for Science Education.

Knapp, M. S. (1997). Between systemic reforms and the mathematics and science classroom: The dynamics of innovation, implementation, and professional learning. *Review of Educational Research, 67*, 227-266.

Lee, O. (1998). *Current conceptions of science achievement in major reform documents and implications for equity and assessment* (Research Monograph No. 12). Madison: University of Wisconsin–Madison, National Institute for Science Education.

Math articles. (1999). *Phi Delta Kappan, 80*, 425-56.

Mintzes, J., Wandersee, J., & Novak, J. (Eds.). (1998). *Teaching science for understanding: A human constructivist view*. San Diego: Academic Press.

Porter, A., Floden, R., Freeman, D., Schmidt, W., & Schwille, J. (1988). Content determinants in elementary school mathematics. In D. A. Grouws & T. J. Cooney (Eds.), *Perspectives on research on effective mathematics teaching* (pp. 96-113). Hillsdale, NJ: Erlbaum.

Resnick, L. B. (1997). The role of teaching and learning in systemic reform. In W. H. Clune, S. B. Millar, S. A. Raizen, N. L. Webb, D. C. Bowcock, E. D. Britton, R. L. Gunther, & R. Mesquita (Eds.). *Research on systemic reform: What have we learned? What do we need to know?* (Workshop Report No. 4, vol. 2, pp. 62-64). Madison: University of Wisconsin–Madison, National Institute for Science Education.

Rodriguez, Alberto, J. (1999). *Equity through Systemic Reform: The Case of Whole-School Mathematics and Science Restructuring in Puerto Rico* (Research Monograph No. 14). Madison: University of Wisconsin–Madison, National Institute for Science Education.

St. John, M., Century, J. R., Tibbits, F., & Heenan, B. (1994). *Reforming elementary science education in urban districts: Reflections on a conference* (Research Monograph). Inverness, CA: Inverness Research Associates.

Smith, M. S., & O'Day, J. (1991). Systemic school reform. In S. H. Fuhrman & B. Malen (Eds.), *The politics of curriculum and testing* (pp. 233-267). Philadelphia: Falmer.

SRI International. (1998). *Introduction to the SSI case studies*. Menlo Park, CA: Author.

Adelman, N. E. (1998). *A case study of Maine's SSI (MAINE: A Community of Discovery), 1992–97*. Menlo Park, CA: SRI International.

Adelman, N. E. (1998). *A case study of Delaware's SSI (Project 21), 1995–97*. Menlo Park, CA: SRI International.

Breckenridge, J. S., & Goldstein, D. (1998). *A case study of Louisiana's SSI (LASIP), 1991–96*. Menlo Park, CA: SRI International.

Goertz, M., & Carver, R. (1998). *A case study of Michigan's SSI (MSSI), 1992–97.* Menlo Park, CA: SRI International.

Goertz, M., Massell, D., & Corcoran, T. (1998). *A case study of Connecticut's SSI (CONNSTRUCT), 1991–96.* Menlo Park, CA: SRI International.

Humphrey, D., & Carver R. (1998). *A case study of New York's SSI (NYSSI), 1993–97.* Menlo Park, CA: SRI International.

Humphrey, D., & Wilson, C. (1998). *A case study of Arkansas' SSI (ARSSI), 1993–97.* Menlo Park, CA: SRI International.

Shields, P., Marsh, J., Marder, C., & Wilson, C. (1998). *A case study of California's SSI (CAMS), 1992–97.* Menlo Park, CA: SRI International.

Zucker, A., & Marder, C. (1998). *A case study of Montana's SSI (SIMMS), 1991–96.* Menlo Park, CA: SRI International.

Standards: Math, science, the NCAA, accountability and school reform [Special issue]. (1998). *Teachers College Record*, 100(1).

Treisman, U. (1997). Response to What have we learned? What do we need to know? In W. H. Clune, S. B. Millar, S. A. Raizen, N. L. Webb, E. D. Britton, D. C. Bowcock, R. L. Gunter, & R. Mesquita (Eds.), *Research on systemic reform: What have we learned? What do we need to know?* (Workshop Report No. 4, vol. 2, pp. 16-21). Madison: University of Wisconsin–Madison, National Institute for Science Education.

Appendix A
Nine Statewide Systemic Initiatives Studied in SRI Case Studies
Synopsized and Ranked According to the NISE Theory
(See Table 1 in text for numerical ratings)

Connecticut

Reform. Independent science, math, and technology (SMT) "Academy" has influence in the department of education, the legislature, most school districts, major professional organizations, and the department of higher education; Academy has affected the curriculum in 19 needy districts, 40 PD providers, the state assessment, and state teacher certification; Academy also has a public relations campaign.

Policy. This state has a "top-down, bottom-up, through-the-middle strategy" of an authoritative, challenging state assessment (no high stakes), plus voluntary aligned program development in schools, districts, and professional organizations; aligned changes in state assessment, teacher certification. The state assessment had been through several cycles of design and modification prior to the SSI, contributing to its quality and authority.

Curriculum. Survey of curriculum in 19 needy districts shows active learning pedagogy, increased enrollment in advanced courses, changes in some district curriculum guidance.

Student achievement. 6-9% more students score proficient on state math assessment, grades 4, 6,8, 1993-97; 7-8% more students score basic/ proficient on NAEP math grades 4,8 (1992-1996). 2-3% more students score proficient on 10th grade state science test over one year (1995-96).

General comment. Policy infrastructure built by Academy appears to have reinforced strong state assessment.

Maine

Reform. The reform agency is an independent "alliance" with links to the Governor, legislature, department of education, higher education, and business groups. It had an impact on curriculum frameworks and assessments and trained a large group of teachers and developed a technical assistance network. The agency probably is sustainable in its reputation and influence.

Policy. Maine established a state assessment with content tests in grades 4, 8, and 11 in 1984 (responding to Nation at Risk). In the 1990s, the SSI worked on alignment of a new set of frameworks, "learning results," and a new version of the state assessment. A group of 7 districts got technical assistance from the SSI and in turn provided technical assistance on a regional basis. Summer "academies" in math and science provided intensive PD. A "leadership consortium" of teachers and others meets to develop common goals and works with the subject matter professional organizations. The combination of these institutions changed SMET educational culture in the state.

Curriculum. About 20% of the state's teachers have received intensive training, while another 40% have received some information and assistance. The training has been evaluated as of high quality and effectiveness. A survey of the classrooms in the technical assistance districts showed high levels of active learning techniques (e.g., 93-100% of elementary teachers emphasizing levels of learning beyond recall; high school classrooms had lower levels, in the 50-75% range). There do not seem to have been any comparisons of reform and nonreform groups or of reform groups over time.

Student outcomes. Maine students showed substantial gains on the state tests of math and science at all tested grade levels in the 1990s (20-65 points on a 300 point scale). Students in assisted schools started and ended this time period about 20 points ahead of the rest of the state.

General comment. Maine's SSI established strong links with all levels of the system (policymakers, delivery infrastructure, schools, and districts), and there were corresponding changes in policy, educational culture, and practice. Students appear to have made strong gains on a state assessment, although the students in assisted schools did not appear to gain more than others.

Montana

Reform. The high school math curriculum reform was led by people active in the national NCTM standards movement, and the MCTM was a leader from the beginning. Awareness of the SSI was high in high schools, as it was known by "practically every math teacher." Two successive governors supported the reform, and the legislature gave three million dollars for a related technology initiative. The curriculum itself was authored by 70 math teachers. The SSI had a public relations arm and published over 600 articles in the media. At its end, the SSI formed an integrated math and science society (partly because of pressure from NSF) and developed an integrated math and science curriculum framework. Still, the absence of science and of the lower grades in math from the reform mission lowers the rating.

Policy. The SSI developed and tested an integrated 4-year high school curriculum (SIMMS) with the first two years intended as the core curriculum for all students. The curriculum was NCTM-like in terms of its vision, topics, requirement of technology, applications, and collaborative learning. Adoption of the curriculum was voluntary. There was no state assessment, but a new accreditation law required the districts to have a curriculum and appropriate assessment, creating a demand for the new curriculum. Within this policy framework, the SSI used consensus building and technical assistance to disseminate the reform. Intensive PD was expected of every math teacher using SIMMS, and workshops were held for thousands of school administrators. State universities contributed overhead on grants to buy computers for teacher preparation. There was a new teacher accreditation requirement; the universities designed new teacher education courses, and state colleges and universities agreed to recognize 3 years of integrated math as meeting the admissions requirement. Again, almost the only weakness is the limitation to high school math, though a different NSF grant supported middle school math. There has also been a decline in state spending and cuts in the Department of education, which the SSI avoided because of its location in a university.

Curriculum. The SIMMS curriculum was used by 40% of math teachers in a majority of the high schools, taken by 25% of the state's high school students, 1/3 of those enrolled in math courses, and ¼ of Native American students. Some professional development was provided for 75% of high school math teachers. Use of the SIMMS curriculum was even higher outside of the academic track where teachers often preferred more traditional courses, especially in the later grades.

Student outcomes. Students in the first two years of the course sequence scored 23 and 14 points higher on a SIMSS open-ended test; students in the third year less so, but these were probably students who previously would not have taken advanced math. Students in the first two years of SIMMS showed no advantage on the PSAT relative to the control group (interpreted to mean that the basic skills levels of SIMMS and non-SIMMS students were equal).

General comment. Montana is a study in contrast between the depth and breadth of its reforms. Looking just at high school math, the strategy was among the most systemic and powerful of all (at least allowing for future scale-up beyond the number of schools already reached). The strategy of developing a new curriculum to meet demand created by a new school certification requirement, plus intensive training of teachers, resulted in rapid adoption of the new courses, especially among those previously not in the academic track. The reform had high visibility in secondary schools, partly because of well-organized professional associations.

Louisiana

Reform. A quasi-independent agency with politically and organizationally skillful leaders from higher education obtained funding from the state boards of higher and K-12 education, had success in getting and coordinating other federal grants; its governance council includes top policy makers; staff includes a full time public relations coordinator; new Governor and reform task force support SSI innovations in frameworks and assessments.

Policy. In first 5 years, 70-75% of resources were spent on high quality, intensive professional development in math and science for 4,100 primarily K-8 teachers (out of about 45,000 teachers in the state); teacher preparation projects in most colleges and universities; new teacher certification requirements. End of first 5 years saw influence on new, aligned frameworks and assessments. End also saw beginning of scale-up efforts through extended PD, school restructuring, and regional assistance to districts. Competency–based curriculum reform and high school exit exam adopted in 1979 are influential, but are not aligned with SSI efforts.

Curriculum. Impact on trained teachers' attitudes was high. Change in classroom practice of trained teachers was broad but uneven in depth.

Student outcomes. Students instructed by SSI teachers scored slightly higher on state (nonaligned) fifth–grade and seventh–grade math tests.

General comment. Judged solely by actual impacts on policy, curriculum, and achievement at the end of the first five years, Louisiana's SSI would have deserved a lower rating. But the reform group has a strong, coordinated influence on policy shown in the recent development of new

aligned frameworks and assessments and new scaling up measures for schools, teachers, and districts.

Michigan

Reform. SSI pushed for alignment of technical assistance with existing strong state assessment and assisted 24 "focus districts" with grants. Technical assistance efforts influenced or produced guidelines for mandatory PD, curriculum and instruction materials on the Web, further alignment of state tests to national standards, advice to regional assistance centers. But staff cut more than 20% by governor at end of SSI.

Policy. State assessment and HS exit exam developed prior to SSI are the leading policy instruments. SSI focused on capacity building through the technical assistance described above, all of which became an infrastructure for reform.

Curriculum. Two–thirds of teachers in focus districts used active learning techniques; all 5 districts visited by evaluation teams had updated their curricula during the SSI period to reflect state assessment; textbook selections in focus districts reflect NCTM standards.

Student achievement. Gains of 5-19 points on state math tests in grades 4,7,11 (but plateau reached around 1995, three years after beginning of SSI); on NAEP math, gains of 6 and 10 points, grades 4 and 9, 1992-96; 7-10 point gains in state science test, grades 5 and 11, in 1996-97 but decline at grade 8; 13% more African-American students proficient on fourth–grade state math test (but gap remains the same); NAEP eighth–grade math gap narrows by 3%.

General comment. Substantial gains in student achievement appear mostly related to earlier policy changes; classroom changes in the focus districts were uneven; and SSI funding was cut at state level, threatening sustainability.

California

Reform. Two teacher networks, in math and science, achieved deep and broad access with teachers, schools, and districts on curriculum and teaching, and each was refunded by NSF under different grants after SSI funding was not renewed. Both networks developed an infrastructure of statewide leadership and regional and school delivery systems.

Policy. California's math and science teacher networks are an interesting example of how a strong "policy" influence can be exerted by a set of intermediate delivery organizations, even when these are no longer supported by state policies. Both networks utilized intensive training sessions in the summer or at other times and had academic year follow-up. Both extended their scale on the basis of what appears to be popular demand (the math network expanding from middle school to elementary school, and the science network adding a math component). The math network (Math Renaissance) used a strategy of curriculum replacement units and influenced curriculum design and textbook selection at the district level. The science network developed a strategy of whole school change and curriculum development at the elementary

school level. Unfortunately, the turmoil surrounding the state policies leaves the future health of the networks somewhat in doubt.

Unfortunately, the good news on the delivery system was matched by bad news in the policies themselves. A back-to-the-basics movement in government policy led to curriculum frameworks being revoked and placed under new development, the state assessment being suspended, the statewide textbook approval and funding becoming less aligned, and the governor pursuing free-standing policy initiatives. Professional reformers in California now have little influence on state policy, but are trying to develop a new consensus.

Curriculum. The breadth and depth of the influence of the networks on the curriculum was strong, based on converging evidence. 38.5 thousand teachers were trained from 2.4 thousand schools in 50% of the state's districts. A study of reform classrooms found change toward standards-based teaching in a majority of classrooms and, in the science classrooms, an average score of 18.75 on a 30-point scale of constructivist teaching. An evaluation found that, in a sample of reform schools, reform-based teaching had achieved sustainable implementation. Districts with reformed schools changed their textbook purchasing to match reform goals.

Student outcomes. In science, students in reform classrooms did not do better than the control group on a specially administered test, but students in schools that had been "under reform" for three years did better than those from schools with two years. In math, a special administration of the new standards exam showed that students from reform schools did better in concepts, skills, and problem solving (with the biggest advantage in skills).

General comment. Based on the effectiveness, power, and scale of its teacher networks, and the systemic policies with which it began the 1990s, California's reform would have deserved a higher rating; but its model systemic policies disintegrated, and the absence of supportive state policy threatened the sustainability of the reform. Also, where they were measured, gains in student achievement were not large, which may be attributable to large declines in financial support for education in the state over many years.

Arkansas

Reform. The SSI was initially supported strongly by Gov. Tucker, but support of the new administration is unclear. Support from departments of both education and higher education. Some aspect of reform reached a large minority of state's teachers and administrators.

Policy. Most resources spent on intensive math and science PD for 35% of all teachers in grades K-4. Trained 22% of all math teachers in grades 5-12 and 22% of science teachers. Also trained 4000 school administrators in leadership academy. Strong state assessment and graduation requirements adopted in 1983 are not well aligned, but SSI is influential in developing new assessment. 3 new levels of SMET teacher certification.

Curriculum. Anecdotal evidence of active learning techniques in classrooms. Science PD effort developed and trained teachers in 17 integrated teaching modules. SSI claims that trained teachers taught 70% of state's students.

Student outcomes. 6-9 point increases in NAEP grade 4 and 8 math scores in the 1990s, increased enrollment in advanced courses, and decrease in students taking remedial education in college probably are mostly caused by basic skills reforms in the 1980s. Student scores were "measurably" higher in schools with 75% or more SSI-trained teachers.

General comment. This state had an extensive PD program that reached many teachers and changed the culture of teaching in the state. But there was limited influence on policy, limited evidence of and probably small impact on classrooms and student outcomes, and lack of clear continuing political support.

Delaware

Reform. The strategy that the SSI began with was judged faulty and heavily revised at the end of five years. The model schools strategy focused on a limited number of schools, lacked a clear vision of goals, produced little change, and was not understood at the district level. The "polished stones" strategy of teachers' developing curriculum units was inefficient and was abandoned in favor of adopting NSF-approved curricula. A new state assessment was suspended after a great many students failed. Summer PD institutes suffered from lack of a means of incorporating school–wide change. But strategies adopted at the end of five years looked more promising (a teacher network built around model curricula, a model of professional development that has been adopted in other states, continuing work on the assessment).

Policy. Few, if any, sustained policy changes were achieved; but the policy profile at the end of 5 years began to look more powerful (especially the combination of curriculum replacement units, teacher networks, and revised professional development).

Curriculum. There was little evidence of curriculum change, and the evaluation found spotty change in a few schools. Participating teachers' attitudes were favorable.

Student outcomes. There was no evidence of a change in student achievement.

General comment. The SRI evaluation overview seems accurate: the Delaware SSI was just acquiring an effective focus at the end of the grant period and could be rated, in a five–stage model of reform developed by the Education Commission of the States (referred to in the SRI case study), as between stages 3 and 4: "transition to a standards-based system, with an emerging infrastructure." The five stages of the ECS model are: (1) non-standards-led system; (2) awareness and exploration of such a system; (3) transition to such a system; (4) emerging new infrastructure to support such a system; (5) predominance of such a system. Note: Under our system of ratings, we probably would classify "awareness and exploration" (ECS stage 2) as a 1, while an "emerging infrastructure" (ECS stage 4), sounds like a 2 under our system; unless there are substantial changes across all four components of reform, policy, classroom, and achievement.

New York

Reform. New York's strategy was to transform 12 urban schools (R & D schools) plus influence state policy. The SSI did pilot new state assessments in the R&D schools, but otherwise had little visibility in state policy. In 1995, in response to NSF, the SSI changed course to emphasize state policy and adopted what appeared to be an unrealistically ambitious plan to transform education in the state.

Policy. Regarding the pilot school strategy, the SSI had difficulty affecting the schools because of complex district bureaucracies, and effects on the districts themselves were minimal. Teachers from these schools attended summer PD institutes, but the institutes were not connected clearly with each other. Regarding state policy, massive cuts were made in the department of education, and reorganization of the department made it more difficult to locate technical assistance. New York had a teaching-oriented school quality review based on British inspectorate, but funding for the program was cut. New assessments and curriculum frameworks were under development, but the SSI had little involvement.

Curriculum. Restructuring progress in the 12 pilot schools was uneven. Only one small elementary school showed deep restructuring. A survey of teachers in the R&D schools showed what appeared to be modest levels of inquiry-based teaching techniques.

Student outcomes. One percent more students in R&D schools reached the proficiency level during third grade on a state math exam (the PEP) than students from other schools. Equivalent gains in the one deeply restructured school were more in the range of 10-20% in both math and science in grades 3 and 6. Science scores were not differentially affected in other R&D schools.

General comment. The state of New York has some promising policies recently developed or under development: new curriculum frameworks, a new assessment aligned with national standards, new rigorous teacher certification. But, in a sense, the SSI chose a "worst of all worlds" strategy: reforming a handful of R&D schools and achieving modest results in that narrow objective, while having little visibility and impact at the state level. It was a good idea to work with urban schools, but the schools and their districts proved difficult to influence, and reform was further impeded by resource deficits at the school level. Professional development was never effectively coupled with the school restructuring strategy. The state department was also rocked by budget cuts.

Appendix B

NISE Protocol Rating for Systemic Reforms

Element/ Component	Breadth (0 to 5)	Depth (0 to 5)
SYSTEMIC REFORM LEADERSHIP AND MANAGEMENT		
A. Vision.	**1,1** What portion of the state system for math education is included in the SSI's vision for reform? **1,2** What portion of the state system for science education is included in the SSI's vision for reform?	**2,1** Does the SSI vision reflect the emphasis of national standards-based reform (i.e., An emphasis on ambitious curricular content for all students)? **2,2** Is the SSI vision coherent; does it address the need to align policies to one another and to new standards? **2,3** Does the SSI vision recognize higher and more equitable student achievement as the ultimate goal of reform? **2,4** Is the SSI vision clear about the kind of teaching and learning that is desired?
B. Strategic planning.	**1,1** What portion of the state system for math education is included in the SSI's strategic plan for reform? **1,2** What portion of the state system for science education is included in the SSI's strategic plan for reform?	**2,1** Is the SSI's strategic plan consistent with the vision for standards-based reform? **2,2** If standards-based reform was underway in the state prior to the SSI, does the strategic plan of the SSI build on what came before? **2,3** Does the SSI's strategic plan include a clear and viable method for scaling up to reach all schools, and math and science classrooms and students in the site? **2,4** Is the SSI's strategic plan specific about **2,4,a** Which organizations and individuals must change? **2,4,b** The type of changes desired of individuals or organizations? **2,4,c** The means by which organizational and individual change will be effected (e.g., Have material and human resources been allocated to encourage changes? Are sanctions and incentives used effectively to promote change)? **2,4,d** Which organizations or individuals will facilitate the change process? **2,4,e** How to monitor the status of the system with respect to desired changes in organizations and individuals? **2,4,f** What constitutes achieving the desired changes in organizations and individuals?

Element/ Component	Breadth (0 to 5)	Depth (0 to 5)
C. Policymaker networking.	**1,1** How broadly does the SSI network with state policymakers? Does the network include the **1,1,a** Governor? **1,1,b** Chief state school officer? **1,1,c** State legislature? **1,1,d** State and local school boards? **1,1,e** Political action groups active in educational area? **1,1,f** Business community?	**2,1** Is standards-based reform the clear focus of interaction between SSI leadership and other state policymakers? **2,2** Does the SSI exert strong influence on policymaker decisions? (i.e., Have policymakers changed existing policy or created new policies in coordination with the SSI and standards-based reform? Alternatively, have policymakers resisted SSI attempts to influence policies that bear positively or negatively on reform?)
D. Professional networking.	**1,1** How broadly does the SSI network with state professional groups? Does the network include **1,1,a** Professional organizations for math and science teachers, and school administrators? **1,1,b** Higher education community for teacher and math and science education **1,1,c** State and local education agencies? **1,1.d** Textbook and test publishers? **1,1,e** Teacher unions?	**2,1** Is standards-based reform the clear focus of interaction between SSI leadership and members of the professional network? **2,2** Does the SSI exert strong influence on other members of the professional network? (i.e., Have organizations representing professional groups made concrete changes to increase coordination with the SSI and standards-based reform? Have these organizations resisted SSI efforts to change organizational practices that undermine reform?)

SYSTEMIC REFORM (cont.)

Element/ Component	Breadth (0 to 5)	Depth (0 to 5)
E. Institutional-ization.	**1,1** To what extent have specific organizations accepted on-going responsibility for the key goals and activities associated with standards-based reform? (i.e., To what extent has the SSI secured organizational homes for key reform activities?)	**2,1** Do the organizations charged with sustaining and expanding the reform have a secure position from which to support reform in coming years? (e.g., Do they have funding, a home, permanent staff?) **2,2** How comprehensively have the organizations identified as responsible for sustaining the reform acted to insure that they are organized internally to value and advocate for reform, and to provide maximum support for the spread and success of reform? For example, have these organizations **2,2,a** Made standards-based systemic reform a prominent part of their own mission? **2,2,b** Staffed up to support reform? **2,2,c** Planned for the maintenance of policymaker and professional networks needed to continue implementation? **2,2,d** Established a system for tracking the progress of reform? **2,2,e** Formulated a strategic plan for advancing reform?
F. Leveraging of resources.	**1,1** How much funding and resources has the SSI leveraged to support standards-based reform? **1,2** From how many sources has the SSI leveraged significant funding and resources to support standards-based reform?	**2,1** Did the SSI tap the funding agents that were obvious candidates to support systemic reform? **2,2** Were the additional resources leveraged by the SSI used in ways consistent with advancing standards-based reform?
G. Public out-reach.	**1,1** To what extent did the SSI involve parents and the public in standards-based reform?	**2,1** Did the public outreach activities of the SSI provide parents and the public with accessible and specific information about standards-based reform/teaching and learning? **2,2** Did the public outreach activities of the SSI seek only to inform parents and the public about standards-based reform, or also seek to actively involve them in some way? (e.g., Did the SSI involve parents in developing new standards, assessments, etc.?) **2,3** Did the SSI monitor or evaluate the effectiveness of public outreach activities?
Subtotal (Average)		

Element/ Component	Breadth (0 to 5)	Depth (0 to 5)
SYSTEMIC POLICY		
	Breadth refers to the extent to which a given component of the site's instructional guidance system affects all relevant actors across the entire state system for math and science education (e.g., Do the standards reach all teachers and students? Does professional development reach all teachers?).	Depth refers to four attributes of the policies associated with a given component of the instructional guidance system.[1] The four attributes include: **2,1** Consistency (Are the site's standards aligned to national standards? Are other components of the instructional guidance system aligned to the site's curriculum standards?)[2] **2,2** Specificity (Is the policy specific enough such that teachers or others will understand what outcomes the policy seeks and what actions the policy requires of them?) **2,3** Authority (Does the policy have the backing of authoritative individuals and organizations, especially ones respected by teachers, administrators, parents, or other key actors?) **2,4** Power (Does the policy include the funding or resources needed for dissemination and implementation? Does the policy entail powerful incentives or sanctions for motivating teachers and others to attend to the policy?)
H. Curriculum standards.	**1,1** Does the site have curriculum standards for all levels of math? **1,2** Does the site have curriculum standards for all levels of science?	**2,1** Are the state standards consistent with national standards? (i.e., Do the state's standards reflect the emphasis of national standards on ambitious curricular content and teaching for understanding?) **2,2** Are the curriculum standards specific enough to provide teachers and others with a clear understanding of what kind of curricula and teaching and learning are desired? **2,3** Are the curriculum standards backed with the authority of individuals or organizations who have legitimacy with teachers, administrators, and the public? (e.g., state and national teacher and administrator professional organizations?) **2,4** Are there powerful sanctions or incentives to encourage teachers to become familiar with the standards and implement them in their classrooms?

[1] The four criteria (i.e., consistency, specificity, authority, and power) used here to characterize and analyze policy tools are developed in Andrew C. Porter and others, "Content Determinants in Elementary School Mathematics," in Douglas A. Grouws and Thomas J. Cooney, Eds., *Perspectives on Research on Effective Mathematics Teaching* (Hillsdale, NJ: Erlbaum, 1988) pp. 96-113. We substitute the term specificity for prescriptiveness, the latter being the term used by Porter and others.

[2] In assessing the *consistency* of the **curriculum standards** developed by a given SSI, the site's curriculum standards will be compared to national standards. In assessing the *consistency* of **other components of the instructional guidance system**, we ask, To what extent is a given component aligned with the site's curriculum standards? This differs from how consistency is conceptualized in the above cited work of Porter and others. Their focus was more on the internal consistency of each piece of the instructional guidance system (e.g., Does a given textbook give teachers consistent signals about what constitutes important curricular content?).

Element/ Component	Breadth (0 to 5)	Depth (0 to 5)
I. Curriculum Frameworks	**1,1** Does the site have curriculum frameworks for all levels of math? **1,2** Does the site have curriculum frameworks for all levels of science?	**2,1** How thoroughly have the curriculum frameworks been aligned to the standards? **2,2** How specific are the frameworks about the kind of curricula and pedagogy that is appropriate to standards-based instruction? **2,3** Do the curriculum frameworks have the backing of authoritative groups (e.g., state professional associations for math and science teachers?) **2,4** Are resources available to produce and disseminate the frameworks so that they are understood and used?
J. Assessment.	**1,1** Does the site have assessments for all levels of math? **1,2** Does the site have assessments for all levels of science?	**2,1** Are the assessments aligned to the site's curriculum standards? **2,2** Are the assessments backed by the state, higher education, teacher professional organizations, and others? **2,3** Are there sanctions or incentives for students or teachers based on performance?
K. Curricular materials.	**1,1** Has the site created new curricular materials, adapted existing materials, or adopted existing materials for all levels of math? **1,2** Has the site created new curricular materials, adapted existing materials, or adopted existing materials for all levels of science?	**2,1** Are the newly created, adapted, or adopted materials aligned to the standards? **2,2** How specific are the materials with respect the kinds of content and pedagogy are appropriate under the standards? **2,3** Are the materials endorsed by the state, teacher professional organizations, and others? **2,4** Are there sanctions or incentives in place to encourage the use of the materials? (e.g., Are the materials provided to the schools free of charge? Are the materials included on the state textbook adoption list?)

SYSTEMIC POLICY (cont.)

Element/ Component	Breadth (0 to 5)	Depth (0 to 5)
L. Equity policy.	**1,1** Does the site address equity considerations in connection with policies pertaining to each component of the instructional guidance system, including **1,1,a** Curriculum Standards? **1,1,b** Curriculum Frameworks? **1,1,c** Assessment? **1,1,d** Curricular Materials? **1,1,e** Teacher preservice/initial licensing? **1,1,f** Teacher recertification? **1,1,g** Teacher/administrator professional development? **1,1,h** Student/ Teacher/ Administrator/ School Accountability? **1,1,i** District/ School capacity-building and improvement?	**2,1** Are policies pertaining to minorities and disadvantaged students clearly grounded in the curriculum standards? **2,2** Are equity policies specific about what constitutes equity in math and science teaching and learning? Do equity policies send clear signals about who is responsible for making the system equitable, and what actions are consistent with equity policy? **2,3** Are there powerful sanctions and incentives in place to add strength to equity policies?
M. Teacher preservice/initial licensing.	**1,1** What portion of the state's preservice teachers of math and science are in programs that are subject to state policy pertaining to preservice program curricula?	**2,1** How comprehensively have the affected programs been aligned to the standards (i.e., How consistent are they with the standards?) **2,2** How specific are the policies that guide preservice/initial licensing regarding practices that support standards-based reform? **2,3** Do policies on teacher preservice/licensing have the backing (i.e., authority) of important groups from higher education (especially the teacher education departments), the state education agency, the state legislature, teacher professional organizations, etc.? **2,4** What resources, or sanctions or incentives have been used to add power to teacher preservice/licensing policies?

Element/ Component	Breadth (0 to 5)	Depth (0 to 5)
N. Teacher recertification.	**1,1** What portion of inservice teachers of math and science are subject to state recertification policies?	**2,1** How comprehensively has recertification been aligned to the standards? **2,2** How specific are the policies for recertification regarding what is appropriate to standards-based reform? **2,3** Do policies on teacher inservice for recertification have the backing (i.e., authority) of important groups from higher education (especially the teacher education departments), the state education agency, the state legislature, teacher professional organizations, teacher unions, etc.? **2,4** What resources, or sanctions or incentives have been used to add power to teacher recertification policies?
O. Teacher/ Administrator professional development.	**1,1** What portion of inservice teachers of math have engaged in professional development during the period of time covered by the SSI? **1,2** What portion of inservice teachers of science have engaged in professional development during the period of time covered by the SSI? **1,3** What portion of inservice administrators have engaged in professional development during the period of time covered by the SSI?	**2,1** How comprehensively has professional development for teachers and administrators been aligned to the standards? **2,2** How specific are the policies for professional development regarding what is appropriate to standards-based reform? **2,3** Do policies on teacher or administrator professional development have the backing (i.e., authority) of important groups from higher education (especially the teacher and administrator education departments), the state education agency, the state legislature, teacher and administrator professional organizations, teacher unions, etc.? **2,4** What resources, or sanctions or incentives have been used to add power to teacher or administrator professional development power to related policies?
P. Student / Teacher / Administrator / School accountability.	**1,1** What portion of each of the following groups is held accountable by the site for its performance in math and science **1,1,a** students? **1,1,b** teachers? **1,1,c** administrators? **1,1,d** schools?	**2,1** How comprehensively has accountability been aligned to the standards? **2,2** How specific are accountability policies as to what constitutes standards-based performance? **2,3** Do policies on accountability have the backing (i.e., authority) of important groups including the state education agency, the state legislature, teacher and administrator professional organizations, teacher unions, etc.? **2,4** Have resources, or sanctions or incentives been used to increase the power of accountability policies?

SYSTEMIC POLICY (cont.)

Element/ Component	Breadth (0 to 5)	Depth (0 to 5)
Q. District / School capacity-building and improvement.	**1,1** What portion of districts in the state are subject to state capacity-building and improvement policies? **1,2** What portion of schools in the state are subject to state capacity-building and improvement policies?	**2,1** How comprehensively have district and school capacity-building activities and improvement efforts been aligned to the standards? **2,2** How specific are the policies for school capacity-building and improvement as to what capacities and improvements are most critical to standards-based reform? **2,3** Do policies on capacity-building and improvement have the backing (i.e., authority) of important groups including the state education agency, the state legislature, teacher and administrator professional organizations, teacher unions, etc.? **2,4** Have resources, or sanctions or incentives been used to increase the power of policies regarding district or school capacity-building and improvement?
Sub Total (Average)		
MATH & SCIENCE STANDARDS-BASED CURRICULUM FOR ALL STUDENTS		
R. Upgraded instruction.	**1,1** Under the SSI, What percentage of the state's students have had access to upgraded instruction for math? **1,2** Under the SSI, What percentage of the state's students have had access to upgraded instruction for science?	**2,1** To what degree is upgraded instruction aligned to the new curriculum standards? (i.e., Does it emphasize teaching for understanding?)
S. School improvement.	**1,1** What portion of the state's schools have engaged in significant school improvement?	**2,1** Under the SSI, how well are school improvement and capacity-building efforts aligned to the standards? (e.g., Is the SSI helping schools establish shared goals and facilitating positive attitudes toward and actions reflective of standards-based teaching and learning?)
T. Equity targeting.	**1,1** What portion of the students in groups that are traditionally under-served in math and science education have been targeted to receive upgraded curricula and instruction under the SSI?	**2,1** To what degree is the upgraded instruction for traditionally under-served students aligned to the new curriculum standards? (i.e., Does it emphasize teaching for understanding?)

Element/ Component	Breadth (0 to 5)	Depth (0 to 5)
U. Data / analysis.	**1,1** For what portion of the state's math classrooms does the site collect data on classroom instructional practices (including delivered curricula and pedagogy)? **1,2** For what portion of the state's science classrooms does the site collect data on classroom instructional practices (including delivered curricula and pedagogy)?	**2,1** To what extent are the data collected on classroom instruction linked to the site's curriculum standards? **2,2** How systematically are data on instructional practice analyzed and reported? **2,3** To what extent do data on instructional practice enter into on-going management of the reform?
Subtotal (Average)		
TOTAL (AVERAGE)		
IMPROVING & EQUITABLE STUDENT OUTCOMES		
V. Value-added to student performance.	**1,1** What portion of the student population is included in the site's measures of math achievement? **1,2** What is the size of any increase in student math academic achievement as indicated by available measures? **1,3** What portion of the student population is included in the site's measures of science achievement? **1,4** What is the size of any increase in student science academic achievement as indicated by available measures?	**2,1** What is the quality of the data the site collects on student performance? (i.e., How complete and reliable are the data?) **2,2** Are the measures included in the data clearly designed to facilitate assessing the effects of standards-based reform on student performance at the site level? **2,3** Do the data include measures that allow for assessing whether or the extent to which gains in student achievement can be associated with or attributed to the SSI or statewide standards-based reform generally?

IMPROVING & EQUITABLE STUDENT OUTCOMES (cont.)

Element/ Component	Breadth (0 to 5)	Depth (0 to 5)
W. Equity of student performance.	**1,1** For what portion of the students in traditionally underperforming groups does the site have data pertaining to changes in performance gaps in math? **1,2** To what degree have math performance gaps between high- and low-achieving groups across the site been reduced under the SSI? **1,3** For what portion of the students in traditionally underperforming groups does the site have data pertaining to changes in performance gaps in science? **1,4** To what degree have science performance gaps between high- and low-achieving groups across the site been reduced under the SSI?	**2,1** What is the quality of the data the site collects on equity of student performance? (i.e., How complete and reliable are the data?) **2,2** Are the measures included in the data clearly designed to facilitate assessing the effects of standards-based reform on the equity of student performance at the site level? **2,3** Do the data include measures that allow for assessing whether or the extent to which increased equity in student achievement can be associated with or attributed to the SSI or statewide standards-based reform generally?
X. Course enrollment/ Attainment/ grades.	**1,1** What portion of the student population is included in the site's measures of student math and science coursetaking, overall attainment, or student performance as reflected by grade distributions? **1,2** What is the size of any increases in student performance as indicated by such measures?	**2,1** What is the quality of the data the site collects on student coursetaking, attainment, and grades? (i.e., How complete and reliable are the data?) **2,2** Are the measures included in the data clearly designed to facilitate assessing the effects of standards-based reform on student performance at the site level? **2,3** Do the data include measures that allow for assessing whether or the extent to which gains in student performance can be associated with or attributed to the SSI or statewide standards-based reform generally?
TOTAL (AVERAGE)		

Single copy price is $7.00. To order copies contact:

CENTER DOCUMENT SERVICE
Wisconsin Center for Education Research
1025 W. Johnson St., Room 242
Madison, WI 53706-1796
608/265-9698

NO PHONE ORDERS. PREPAYMENT REQUIRED FOR ORDERS UNDER $20.00.

Price is subject to change without notice.

National Institute for Science Education
University of Wisconsin–Madison
1025 West Johnson Street
Madison, WI 53706

(608) 263-9250
(608) 262-7428 fax
niseinfo@macc.wisc.edu
http://www.wcer.wisc.edu/nise